Business Register Profile: U.S. Bureau of the Census [1]

INTRODUCTION

1.   This paper presents a profile of the Census Bureau's Business Register (BR). It gives an overview of the BR, covering basic topics such as purpose, scope, statistical units, information content, and data sources. Further, it highlights selected characteristics of the business population, as the register represents it, to illustrate key aspects of BR composition and structure. Lastly, it outlines BR maintenance procedures.

BR OVERVIEW

2.   The BR is a central, multipurpose, business list that the Census Bureau has maintained continuously since 1972. Established under authority of Title 13, United States Code, and a 1968 designation by the Office of Management and Budget, this statistical resource is a central element in the infrastructure of the Census Bureau's economic programs. An annual Congressional appropriation provides BR funding under the Census Bureau's "General Economic Programs" line item.

Confidentiality of BR Information

3.   All BR information about individual businesses is confidential under Title 13, Unites States Code. Additionally, Federal tax information incorporated in the BR is confidential under both Title 13 and the Internal Revenue Code (Title 26), whereas other administrative data in the BR are confidential under administrative directives and contractual agreements with supplying agencies. Therefore, BR information about individual businesses is available only for internal use by authorized[2] Census Bureau programs, and the agency cannot supply business lists or other BR information about the characteristics of specific business units to external parties—not even to other organizations within the Federal government.

Purpose

4.   The BR's primary purpose is to serve as a comprehensive source of list frames for authorized[2] economic censuses and surveys. In this capacity, it identifies the set of statistical units that represents an economic data collection's target population, which is defined by a specific reference period and scope (type of statistical unit; ownership characteristics; industrial, geographic, and size classifications; and other factors). Critical to this purpose is the BR's ability to provide a coherent methodological framework; complete, unduplicated, and up-to-date coverage; consistent industrial, geographic, and size classifications; accurate contact information; and effective support for data collection and processing operations.

---

[1] This paper reports the general results of analysis undertaken by Census Bureau staff. The views expressed are attributable to the author and do not necessarily reflect those of the Census Bureau.

[2] Tax information in the BR may be used only for certain programs—those authorized under Chapter 5 of Title 13. This includes the economic census and most periodic surveys on subjects covered by the census.

5.   Secondarily, the BR serves as all of the following, listed roughly in order of importance:

  •   A central repository for administrative records information (mostly Federal tax data), used by authorized[3] economic programs.

  •   A central support facility for collection and processing.

  •   The source of basic employment and payroll measures summarized by industry and geographic area in the annual County Business Patterns and ZIP Business Patterns statistical series.

  •   A research resource for studies performed by the Census Bureau's Center for Economic Studies and other authorized investigators.

  •   A ready-made data source for custom tabulations and other special studies (reimbursable projects).

  •   The basis for longitudinal studies that track units through reorganizations or changes in ownership and provide information on business demographics.

Scope

6.   The scope of the BR includes economic entities defined by the following characteristics:

6.1   **Legal Status**—The BR covers entities that operate legally (i.e., entities that are identified by the Master File systems of the Internal Revenue Service [IRS]). The BR and the Census Bureau's economic programs make no attempt to account for underground (illicit) economic activity.

6.2   **Geographic Classification**—The BR covers entities that operate within the United States and its territories (Guam, the Northern Mariana Islands, Puerto Rico, and the United States Virgin Islands).

6.3   **Employer Status**—From its inception through 1993, the scope of the BR was limited to entities with paid employees, except for years covered by the economic censuses (those ending in '2' and '7'), when coverage included nonemployer entities (businesses without paid employees) that were subject to Federal income tax. Examples of nonemployer entities include self-employed persons and other businesses operated exclusively by working proprietors or partners. Since 1994, an annex to the BR has provided annual coverage for nonemployer entities that are subject to Federal income tax. This paper focuses almost exclusively on the part of the BR for entities with paid employees.

6.4   **Industrial and Ownership Classifications**—The industrial scope of the BR depends partly on private versus government ownership or control.

  •   **Privately-Owned Entities**—For privately-owned entities, the BR systematically covers all industries except employers classified as private households (North American Industry Classification System [NAICS] 814). Taxpaying entities classified in other industries that are out of scope to the Census Bureau's economic programs are present on the BR but generally are not resolved into establishment units unless they are part of a multi-establishment enterprise engaged in other activities that are in-scope to the Census Bureau's economic programs. The out-of-scope industries are: agriculture, forestry, fishing and hunting (NAICS 11); rail transportation (NAICS 482); the U.S. Postal Service (part of NAICS 491); certificated passenger air carriers (part of NAICS 481111); elementary and secondary schools (NAICS 6111); colleges and universities (NAICS 6113); labor organizations (NAICS 81393); political organizations (NAICS 81394); religious organizations (NAICS 8131); and public administration (NAICS 92; however, the Census Bureau maintains a separate register of government units, known as the Governments Integrated Directory).

---

[3] Tax information in the BR may be used only for programs authorized under Chapter 5 of Title 13.

- **Government-Owned/Controlled Entities**—For government-owned/controlled entities, the BR covers only selected industries that are in-scope to the economic censuses; they are: beer, wine, and distilled alcoholic beverages wholesalers (NAICS 4228); beer, wine and liquor stores (NAICS 4453); monetary authorities—central bank (NAICS 521); Federal and Federally-sponsored secondary market financing (part of NAICS 522294); and hospitals (NAICS 622).

Statistical Unit Definitions

7.   The BR explicitly represents three basic statistical unit types, defined as follows:

7.1   **Establishment**—An establishment is an economic unit, generally at a single physical location, where business is conducted or where services or industrial operations are performed. Examples include a mine, factory, warehouse, sales office, grocery store, bank, hotel, movie theater, doctor's office, museum, and central administrative office. The establishment is the BR's most basic statistical unit; it serves as the primary collection unit for the economic censuses and some economic surveys (for example, the Annual Survey of Manufactures); also, it may provide building blocks for constructing more aggregated collection units for other economic surveys. Further, the establishment is the fundamental unit to which industrial classifications apply.

There are three minor *exceptions* to the general establishment definition, as implemented by the BR and the Census Bureau's economic programs:

- **Distinct Activities as One Location**—In cases where distinct and separate economic activities are performed at a single physical location, each activity may be treated as a separate establishment, provided that: (i) no one industrial classification includes those combined activities, (ii) the employment in each activity is significant, and (iii) the organization can prepare a separate report on basic economic measures for each activity. A central administrative office at the location of a retail chain's "flagship" store would be a common example.

- **Consolidated Enterprise Units**—As a practical matter, the BR may not resolve as separate establishments two or more locations operating under common ownership or control if: (i) all such locations would be classified in industries that are out of scope to the Census Bureau's economic programs (for example, railroads), or (ii) the locations are classified in the same industry and geographic area and combined employment for all such locations is small, generally fewer than 10 persons.

- **Field Sales Force "Establishments"**—In cases where a business' field sales force employees cannot be associated with a traditional establishment (single physical location), those employees are grouped by state and assigned to special statewide units. Statewide units are treated as establishments in most respects; however, their geographic classifications (see Information Content, below) localize them only to state and do not place them in any smaller geographic area (county or place).

7.2   *EIN Entity*—An EIN (Employer Identification Number) entity is an administrative unit that the IRS has assigned a unique identifier for use in tax reporting. In practice, this identifier is used more broadly by administrative records systems (for example, administrative records of the Federal/State Unemployment Insurance [UI] systems include the employer's EIN). Therefore, its presence on the BR is critical for linking administrative records information to related establishment and enterprise units. The EIN entity also is used as a sampling/collection unit in some surveys (notably, annual and monthly surveys covering a broad range of services).

7.3   *Enterprise*—An enterprise is an **economic** unit comprising one or more establishments under common ownership or control. In the BR's implementation, this unit represents the top-level U.S. parent company and all subsidiary U.S. companies of which the parent owns or controls a majority (more than 50 percent) interest plus all EINs and establishments affiliated with the parent and its subsidiaries.

Statistical Unit Relationships

8.   The relationships among the BR's statistical units constitute a simplified (abstracted) model of the real-world relationships among business/organizational units, administrative records, and the Census Bureau's statistical collections. Designed some 30 years ago, this model may fail to capture the full subtlety and complexity of more intricate organizational structures that sometimes characterize today's enterprises. Similarly, it may not fully reflect the administrative and statistical reporting practices that have evolved since the BR was established.

9.   The basic relationships among the BR's statistical units are as follows:

9.1   *An* **Enterprise**:

- Owns or controls at least one and perhaps many establishments. While the enterprise comprises one or more legal entities (natural or legal persons), those legal entities are not represented on the BR as a distinct statistical unit type. Similarly, the enterprise may direct its establishments through subsidiary companies or other intermediate management structures (for example, divisions or regions), neither of which are represented on the BR as distinct statistical unit types (however, any subsidiary, divisional, or regional headquarters office would be represented on the BR as an establishment engaged in management of companies and enterprises (NAICS 551).

- Files administrative reports (for example, those related to payroll and income taxes and to unemployment insurance) with governing authorities for at least one and perhaps more EIN entities.

9.2   An *EIN* **Entity**:

- Supplies administrative records information about one or more establishments.

- Is associated with one and only one enterprise. While EIN entities often bear a close correspondence to the enterprise's legal entity(ies), no one-to-one association is guaranteed, as a legal entity may have more than one EIN.

9.3   An **Establishment**:

- Is affiliated with one and only one enterprise at a given time.

- Is covered by the administrative reports of at least one and perhaps many EIN entities.

Single- Versus Multi-Establishment Enterprises

10.   There is great variation in the complexity of business organizations. A most basic and useful distinction along the complexity dimension is one between single- and multi-establishment enterprises. As subsequent discussion will show, this distinction has important implications for maintaining the BR through the use of administrative records information.

- **Single-Establishment Enterprises**—In the most elementary case, the enterprise, EIN entity, and establishment are one and the same. This kind of business organization is known as a single-establishment enterprise or, more concisely, a single unit. The simplicity of such a business makes it's representation on the BR easy to maintain by using administrative records as the principal (nearly exclusive) source of information for updates (see BR Updates, below, for more information).

- **Multi-Establishment Enterprises**—In more complex cases, the enterprise has one EIN entity or more and two establishments or more. This kind of business organization is known as a multi-establishment enterprise or, more concisely, a multiunit. The typical multiunit remains relatively small and simple, generally having just one EIN entity and two or three establishments; however, a few of the United States' largest and most complex enterprises have thousands of EIN entities and more than ten thousand establishments.

11.	As Table 1 (below) shows, multiunits amount to only 3 percent of the BR's employer enterprises; however, they represent a disproportionate share of economic activity as gauged by other indications. For example, they account for 6 percent of the BR's EIN entities, 23 percent of its employer establishments, and 62 percent of its total employment for the pay period including March 12 (full- and part-time).

Type of Enterprise	Number of— [Percent of Column Total]			
	Active Enterprises	Active EIN Entities	Active Establishments	Employees (Mid-March)
Single-establishment	5.0 million [97%]	5.0 million [94%]	5.0 million [77%]	36.8 million [38%]
Multi-establishment	0.2 million [3%]	0.4 million [6%]	1.5 million [23%]	59.9 million [62%]
Total	5.2 million [100%]	5.4 million [100%]	6.5 million [100%]	96.7 million [100%]

Table 1. Enterprise Composition of the BR [enterprises with paid employees]

Information Content

12.	The BR records several dozen attributes for each statistical unit and organizational information that links associated establishments and EIN entities within an enterprise. Some of the information in the BR directly supports statistical program uses, including survey frame construction and access to administrative data, whereas other information is used by BR systems to perform internal management and auditing functions. The following categories of BR information are of general interest for statistical program uses:

12.1	**Statistical Unit Identifiers:**

- **Census File Number (CFN)**—A unique 10-digit CFN serves as the primary identifier for each establishment and designates current ownership. In the case of multiunit establishments, the first 6 digits of the CFN identifies the parent enterprise (includes the parent's Alpha; see below).

- **Alpha**—A unique 6-digit Alpha serves as the primary identifier for each multiunit enterprise and is embedded in the CFN of each affiliated establishment.

- **Permanent Plant Number** (PPN)—A 10-digit PPN identifies establishments without regard to current ownership, and its continuity is maintained through reorganizations and changes in ownership.

- **EIN**—The IRS-assigned, 9-digit EIN uniquely identifies EIN entities. BR affiliation information links the EIN to both the parent enterprise and to the establishments that it covers.

12.2	**Business/Organization Name**—The BR records both primary and secondary names, permitting identification of both legal and trade designations.

12.3	**Other Contact Information**—The BR's contact information includes:

- **Mailing Address**—Components are street address or post office box, post office (city or town) name, postal state abbreviation, and ZIP Code. Mailing address is available only for single units and EIN entities; the parent enterprise's address generally serves as the mailing address for multiunit establishments.

- **Physical Location Address**—Components include street address, post office (city or town) name, postal state abbreviation, and ZIP Code.

- **Attention information**—information for directing inquiries within business organizations (e.g., "Attn: Government Reports Office").

- **Primary reporting official**—the person's name, telephone number, and FAX number (generally, the person responsible for reporting to the economic censuses and/or the annual Company Organization Survey).

12.4 **Industrial Classification**—The BR classifies each establishment according to its primary activity. Additionally, the BR classifies each multiunit enterprise according to its predominant activity, defined as the industry group that accounts for the largest share of employment. However, the BR's classification of multiunit enterprises is used only for certain sampling operations and not for statistical presentation.

The BR includes classification codes for both of the taxonomies identified below, but only the second is fully implemented at this time (we introduced the NAICS coding scheme for the 1997 Economic Census and are now are implementing it fully in a redesigned BR database). Further, the BR includes indicators that show the source (e.g., Census Bureau collection, administrative records), reliability, and year of last update for the current classification.

- **North American Industry Classification System (NAICS) 1997**—The 1997 Economic Censuses marked the Census Bureau's first use of the newly developed NAICS taxonomy. The BR incorporates the resulting census-based classifications as an 8-digit code that represents the 5-digit NAICS industry suffixed by a 1-digit U.S. industry extension and the Census Bureau's 2-digit subindustry extension.[4]

- Standard **Industrial Classification (SIC), 1987 Revision**—The BR records a 6-digit code that represents the 4-digit SIC industry suffixed by the Census Bureau's 2-digit subindustry extension.[4]

12.5 **Type of Operation Classification**—These classifications supplement industrial classifications by distinguishing basic operating modes, such as the following:

- For wholesale trade—Type-of-operation categories distinguish wholesale merchants (establishments selling primarily on their own account); manufacturers' sales branches and offices (owned by multiunit manufacturing enterprises); and wholesale agents, brokers, and commission merchants (selling primarily on a commission basis).

- For auxiliary establishments—Type-of-operation categories distinguish multiunit establishments engaged primarily in performing management and support services for other establishments of their parent enterprise.

12.6 **Geographic Classification**—Automated geographic coding systems classify every establishment with respect to each of the geographic area types listed below, based on physical location address (ZIP Code; postal state abbreviation; post office [city or town] name; and street address (the last applies mostly in urbanized areas with city-style addresses). Additionally, **metropolitan statistical are**a and **consolidated metropolitan statistical area** classifications can be derived from the combination of State and County code (State and Place in New England, were towns are treated as primary political subdivisions).

- **States**—This classification identifies the 50 states, the District of Columbia, and other statistically equivalent areas, including Guam, the Northern Mariana Islands, Puerto Rico, the United States. Virgin Islands, Nationwide, and Ships at Sea (a total of 57 such areas).

- **Counties**—This classification identifies 3,284 counties (parishes in Louisiana) and statistically equivalent areas, including the District of Columbia, independent cities, and Alaska's boroughs/census areas. Counties (parishes) are the major political subdivisions of most states.

- **Places**—This classification identifies 7,671 places for which the economic censuses compile statistics (generally, cities or towns having 2,500 inhabitants or more); additionally, it

---

[4]   Subindustry extensions often are '00', which indicates that the Census Bureau has defined no subindustry detail.

identifies a balance-of-county area for all counties and statistically equivalent areas that are not coextensive with or wholly contained by separately identified places.

- **Census Tracts and Block Numbering Areas**—This classification identifies 62,276 statistical subdivisions of selected counties that are defined by a local committee of data users and are used primarily in the United States' 10-year population census. Census Tracts and Block Numbering Areas are not used for presentation of economic statistics; rather, they serve as building blocks for larger geographic areas.

- **Census Blocks**—This classification identifies 7,017,427 statistical subdivisions of Census Tracts and Block Numbering Areas, generally bounded by physical features (e.g., streets or streams) and the boundaries of larger legal or statistical areas; again, these areas are used primarily in the population census. Census Blocks are not used for presentation of economic statistics; rather, they serve as building blocks for larger geographic areas.

12.7 **Legal Form Classification**—Each statistical unit is classified as one of the following, based primarily on the kind of business income tax form filed:

- Sole proprietorship

- Partnership

- Corporation

- Cooperative association

- Government-owned or controlled entity (covered only in selected industries).

- Other (for example, trust, foundation, estate)

12.8 **Federal Income Tax (Taxable/Tax-Exempt) Status**— Each statistical unit is classified in one of the categories shown below. This distinction affects collection and presentation of the basic dollar volume measure, which is value of shipments, sales, or receipts for taxable units and revenue for tax-exempt units; additionally, the latter use operating expenses as a secondary dollar volume measure.

- Subject to Federal income taxes (generally profit-making).

- Exempt from Federal income taxes under Section 501, Internal Revenue Code (non-profit).

12.9 **Basic Measures of Size and Economic Activity**—The BR records actual values for all measures listed below plus size class codes for the employment variable:

- **Number of Employees for the Pay Period Including March 1**2— The BR employment measure includes full- and part-time employees, salaried corporate officers and executives, and persons on paid absences (for example, vacation or sick leave); it excludes working proprietors and partners, volunteers, other persons in non-pay status, pensioners, persons working as independent contractors, and persons supplied by employee leasing (co-employment) or temporary help services. The BR records the employment of establishments and EIN entities for the current year and two prior years, and it imputes missing values using regression-based estimates that take seasonality into account. An associated indicator identifies the source of the employment value, including imputation. Additionally, the BR separately records current-year employment totals for enterprises, computed as the sum of all establishments and the sum of currently active establishments.

- **First Quarter Payroll**—BR payroll consists of wages, salaries, bonuses, commissions, and other compensation paid to employees, including employee contributions to qualified pension plans and amounts paid to corporate officers and executives; it excludes amounts withdrawn from profits by proprietors and partners. The BR records the first quarter payroll of establishments and EIN entities in units of $1,000 for the current year and two prior years; 2nd, 3rd, and 4th quarters also are available for single units and EIN entities. The BR

imputes missing values using regression-based estimates that take seasonality into account, and an associated indicator identifies data source, including imputation.

- **Annual Payroll**—BR payroll is defined above, under first quarter payroll. The BR explicitly records annual payroll of multiunit establishments in units of $1,000 and imputes missing values using regression-based estimates that take seasonality into account. An associated indicator identifies data source, including imputation. Additionally, an annual payroll figure may be derived as the sum of four quarters for single units and EIN entities.

- **Value of Shipments, Sales, Receipts, or Revenue**— BR output measures comprise a set of four separate figures, the specific content of which varies somewhat, depending on the type of income tax return that the business/organization files. These figures may include total revenue, gross income, operating receipts (gross receipts or sales less returns and allowances), interest income, and gross rents. The BR records these items in units of $1,000 only for single units and EIN entities and only as available (the BR has incorporated yearly output measures only since 1994, and there presently is no imputation of missing values). Statistical applications may derive a single output measure by combining components, as appropriate (for example, interest income would be included only for certain financial industries, and gross rents, only for certain real estate industries).

- **Assets**—The BR records this measure in units of $1,000 only for single units and EIN entities and only as available (the BR has incorporated yearly assets measures only since 1994, and there is no imputation of missing values).

12.10 **Basic Data from the 1997 Economic Census**—Content includes 1997 census data for NAICS, SIC, type-of-operation, and geographic codes; value of shipments, sales, receipts, or revenue; annual payroll; first quarter payroll; number of employees for pay period including March 12; and selected other items that may vary by sector.

12.11 **Basic data from the 1992 Economic Census**—Content includes 1992 census data for SIC, type-of-operation, and geographic codes; value of shipments, sales, receipts, or revenue; annual payroll; first quarter payroll; number of employees for pay period including March 12; and selected other items that may vary by sector.

12.12 **Other Status, Coverage, and Control Information**—BR content includes indicators that show whether a unit is currently active, idle/inactive but still owned, closed, or sold/leased to another operator; other coverage control codes; dates associated with recent coverage changes; and if applicable, the identity of predecessor and/or successor units.

Data Sources

13. The BR integrates information from several sources to achieve a practical and effective balance among competing demands for comprehensive coverage, diverse and accurate content, timely updates, low cost, and minimal response burden. The data sources and their respective roles in BR construction and maintenance are described below:

13.1 **Administrative Records**—Administrative records are the foundation of the BR. They provide indispensable information that is low in cost, timely, comprehensive, and generally quite accurate. Further, Title 13 requires the Census Bureau to use administrative records in lieu of direct collection whenever practical and feasible. They satisfy much of the BR's substantial data requirement while imposing minimal response burden. The BR's principal administrative records suppliers are as follows:

- **Internal Revenue Service (IRS)**—The Internal Revenue Code (Title 26, United States Code), authorizes the IRS to disclose tax information to the Census Bureau "for the purpose of, but only to the extent necessary in, the structuring of censuses…and conducting related statistical activities authorized by law." Related IRS regulations stipulate specific content, uses, and further conditions on the disclosure of this tax information. Both the statute and the

regulations require the Census Bureau to safeguard the confidentiality of tax information. This information comes from the following IRS sources:

Business Master File (BMF)—The BMF identifies EIN entities that represent business, organizational, and agricultural taxpayers known to the IRS; sole proprietorships *without* paid employees are excluded and are identified exclusively from business income tax returns (see below). The Census Bureau receives an extract of BMF information for all EINs during May of each year, with a volume of 24 million records annually. Further, the agency receives monthly BMF extracts containing updates for recent additions, deletions, and other changes, with a total volume of 12 million records annually. Content of these BMF extracts includes EIN; proprietor's Social Security Number (SSN), if applicable; and other identifying information. Further, it provides legal and trade names; mailing and physical location addresses; principal business activity (industrial) classification; and selected control, status, and processing indicators. BMF information is critical to the BR, particularly for identifying newly established EIN entities that represent business births.

Payroll Tax Returns—Business and organizational employers file the Employer's Quarterly Federal Tax Return, IRS Form 941 series, which is of primary importance; agricultural employers file the Employer's Annual Tax Return for Agricultural Employees, Form 943 series. The Census Bureau receives weekly files from current IRS processing of both forms, with total volume amounting to 22 million records annually (about 5.5 million Form 941 records per quarter). Both types of return identify taxpayers by EIN, provide total employment for the pay period including March 12, and indicate the tax period covered. Additionally, Form 941 provides data by quarter and Form 943 by calendar year for three wage measures: total compensation, social security wages and tips, and Medicare wages. Of these wage measures, the Medicare figure generally is closest to the BR payroll definition, because it includes employee contributions to qualified pension plans (total compensation excludes this component) and because it is not subject to an annual reporting ceiling (social security wages and tips have an annual cap for each employee).

The Census Bureau receives reasonably timely information from quarterly payroll tax returns, the IRS Form 941 series. The deadline for filing this return is one month after the end of the reference quarter, and that date corresponds closely to a sharp rise in the flow of records for the period. The weekly flow peaks just 9 weeks after the end of the quarter or 5 weeks after the filing deadline, suggesting that the typical lag due to IRS processing is only a few weeks. Cumulative receipts exceed the 50 percent mark just 9 weeks after the end of the quarter, 90 percent at 11 weeks, and 99 percent at 36 weeks. After that, a trickle of returns (a few hundred per week, presumably late filers or amended returns) continues for several months, but they do not add substantially to total volume.

Business Income Tax Returns—Sole proprietors are identified by Social Security Number (SSN) and file a U.S. Individual Income Tax Return, IRS Form 1040, supplemented by Schedule C for Profit or Loss From Business and/or by Schedule E for Self-Employment Income. Other business and organizational taxpayers file one of the following, depending on their legal form of organization and Federal income tax status: U.S. Partnership Return of Income, Form 1065; Corporation Income Tax Return, Form 1120 series; or (selectively, depending on receipts and assets size) Return of Organization Exempt from Income Tax, Form 990 series. Lastly, parent corporations that file a consolidated income tax return for themselves and qualifying subsidiaries also submit an Affiliations Schedule, Form 851, that identifies by EIN the parent and those subsidiaries covered by the consolidated return. These annual business income tax returns provide basic measures of business receipts or revenue (specific content varies by tax form; see Data Content, above) and assets; most returns also provide a principal business activity (industrial) classification. The Census Bureau receives weekly files from February through December, which contain data from current IRS processing.

The flow of business income tax these records is somewhat more spread out and less timely than is the case with quarterly payroll tax returns. Still, there is a sharp and steady

increase in weekly receipts that peaks just 5 weeks after the March 15 filing deadline for corporations with a fiscal year that corresponds to the calendar year, partnerships, and exempt organizations and the April 15 filing deadline for sole proprietorships. Another substantial peak occurs approximately six months later when an available extension of the filing deadline expires. Cumulative record counts top 50 percent of total approximately 21 weeks after the end of the reference period, 90 percent at 36 weeks, and 99 percent at 47 weeks.

- **Social Security Administration (SSA)**—New business and organizational taxpayers (i.e., births) file an Application for Employer Identification Number, Form SS-4, with the IRS. The IRS forwards those forms to the SSA, which assigns an industrial classification to the applicant based on a principal activity description and other information reported on the form. The Internal Revenue Code and its regulations authorize the Census Bureau to receive information from Form SS-4, and the IRS permits the SSA to release the data directly.

  Form SS-4 content supplied to the Census Bureau includes EIN, NAICS code (4-digit), geographic information, estimated employment, and other classification/status indicators. The Census Bureau receives monthly files from current SSA processing, which lags Form SS-4 filing by some 3 months. Total volume amounts to 1.2 million records annually.

- **Bureau of Labor Statistics (BLS)**—The BLS maintains a separate business register, known as the Business Establishment List (BEL), based on information collected to administer cooperative Federal/state unemployment insurance programs. State employment security agencies supply industrial classifications and other data about private- and public-sector employers to the BLS as part of the Covered Employment and Wages (ES-202) program. Under authority of an Office of Management and Budget directive and a memorandum of understanding between the agencies, the BLS shares certain industrial classifications with the Census Bureau. Each quarter, the Census Bureau prepares a file of EINs that identify BR establishments for which a BLS classification is requested. The BLS refers each of these EINs to their BEL and returns the corresponding NAICS code whenever one is found. [Note: This operation was suspended in February 2000; we hope to resume it early in 2002.]

13.2   Census Bureau Collections

- **Company Organization Survey (COS)**—The COS is a register proving or profiling survey, done specifically for the purpose of maintaining BR information about the establishment composition, organizational structure, and operating characteristics of multi-establishment enterprises. A separate collection for this purpose is necessary because administrative records do not resolve the relationships among multiunit enterprises, their EIN entities, and their establishments, as the BR requires.

  The Census Bureau conducts the COS annually. In years covered by the economic censuses (those ending in '2' and '7'), census and COS programs are combined and the latter collection is limited to basic company affiliation information and to coverage of establishments that are outside the industry scope of the census; meanwhile, the census collects information about all in-scope multiunit establishments. In all other survey years, the COS directs inquiries to approximately 55,000 multi-establishment enterprises, which operate more than 1.1 million establishments. This annual survey panel is drawn from the BR population of nearly 185,000 multi-establishment enterprises, which operate 1.6 million establishments. The procedure for constructing this panel selectively targets enterprises that are most likely to report changes in establishment composition, organizational structure, and/or operating characteristics, based on enterprise size and complexity and on administrative records indications (determined by applying selection rules to associated EIN entities). Additionally, the panel includes a small probability sample of enterprises not selected by the targeting procedure. Enterprises included in each year's panel account for approximately 80 percent of multiunit employment and payroll.

The COS is a combination mail/electronic collection. More than 600 larger enterprises, which account for about 30 percent of COS establishment coverage, return their reports by automated/electronic means (mostly by Computer Self-Administered Questionnaire). All other respondents return a paper questionnaire. Collection operations begin late in December of the reference year and continue through mid-August of the following year.

Data content is identical for all reporting modes. The instrument includes inquiries on ownership or control by a domestic parent, ownership or control by a foreign parent, and ownership of foreign affiliates. Further, the instrument lists an inventory of establishments belonging to the enterprise and its subsidiaries, and it requests updates to the inventory, including additions, deletions, and changes to information on each establishment's EIN, name and address, and industrial classification. Finally, it collects each establishment's end-of-year operating status, employment for the pay period including March 12, first quarter payroll, and annual payroll. These COS inquiries, combined with economic census inquiries during years covered by the census, are the primary source of the information that the BR records for multiunit establishments; therefore the COS and economic census definitions for employment, payroll, and similar content are identical to the BR definitions given above.

- **Economic Censuses**—The economic census, done at 5-year intervals (covering years ending in '2' and '7'), is a comprehensive enumeration of the United States business population and, therefore, a valuable source of information for BR maintenance. Of particular importance are identification of new multiunits and other coverage improvements resulting from systematic analysis of census data, updated address information, and more accurate industrial classifications based on detailed census collections for value of product and/or service outputs by category and other classification factors. Economic census and COS programs are closely integrated, ensuring timely BR updates based on results of these collections.

- **Current Economic Surveys**—Although administrative records, the COS, and the economic censuses provide the great majority of the information needed to maintain the BR, the Census Bureau's monthly, quarterly, and annual surveys supply additional updates. For example, the Annual Survey of Manufactures (ASM) is closely integrated with the COS and economic census, and it provides valuable feedback of coverage and classification information for manufacturing enterprises and their establishments. Similarly, the monthly economic surveys often are the first to identify new multiunit establishments, changes in ownership, and updated address information, and they feed that information back to the BR as well.

BR Updates

14.   The Census Bureau maintains its business register as a relational database with several separate tables for each of the BR's three basic statistical units. Additionally, the database has relational structures that represent organizational linkages among these statistical units. Procedures for updating this database vary with respect to mode, data source, type of unit, content, and frequency. Table 2, below, presents a high-level summary of these procedures, and the discussion that follows expands somewhat on the subject.

14.1   **Mode**—BR maintenance processes use two basic updating modes:

- **Interactive Transactions**—Interactive transactions record changes initiated one-at-a-time by BR specialists and survey analysts based primarily on information from the COS, the economic censuses, and other economic collections. These changes tend to be relatively complex, as they frequently affect critical information about multiunit composition and organization. Further, they often require considerable knowledge about the BR and business organization and depend upon expert judgement in the evaluation of variable and sometimes inconclusive source information.

- **Batch Updates**—Batch updates apply changes to many thousands of statistical units in a single processing run based on data from standard input files. By necessity, these high-volume changes are relatively simple. They may add new EIN entities (births); post employment, payroll, or receipts measures from the latest administrative records files; or provide improved data for attributes such as industrial classification or address; however, they generally have no direct effect on critical information about multiunit composition and organization.

14.2 **Variation by Type Statistical Unit**—BR maintenance procedures and the data sources they use vary, depending on type of statistical unit. As a general rule, BR updates rely almost exclusively on administrative data applied by highly automated batch processes when a statistical unit's organizational relationships are very simple and it's links to administrative records are direct. On the other hand, BR updates become increasingly dependent on Census Bureau collections and require more human intervention through interactive updates when a unit's organizational relationships are complex or it's links to administrative records are indirect. As a useful approximation, this difference in BR data sources and maintenance procedures comes down to a distinction between single units, which depend mostly on automated updates, and multiunits, which require much more human intervention.

- **Single** Units—Single units are a special case where the enterprise, the EIN entity, and the establishment generally are one and the same. Further, the single unit's EIN provides a one-to-one link to administrative records of the IRS, the SSA, and the BLS. Because of these special circumstances, administrative records often provide all the information that the BR requires for single units. Further, maintenance procedures that rely on administrative records impose no additional response burden, are very low in cost, and usually equal or exceed the timeliness achieved by separate register proving surveys. As a result of these advantages, the Census Bureau uses administrative records to the greatest possible extent for maintaining BR information about single units. This approach applies broadly, since the BR records more than 5 million single units with paid employees. Additionally, the BR's nonemployer annex, which records another 14 million single units without paid employees, is derived almost exclusively from administrative records.

  Single units generally require more costly interactive updates only for selected changes based on Census Bureau collections and data evaluation; examples include assignment of improved industrial classifications and the maintenance of linkages between predecessor and successor units when the enterprise has a reorganization or change in ownership. Interactive updates for single units in the BR's nonemployer annex are extremely rare.

- **Multiunits**—The relationships among EIN (administrative records) entities, multiunit enterprises, and multiunit establishments vary from case to case. Because of this variation, administrative records are not sufficient to maintain BR information for multiunits, and the Census Bureau must undertake its own collections, primarily the COS and the economic censuses, for this purpose. Further, BR maintenance tasks that affect multiunit composition and organization usually have a level of complexity that requires intervention by a BR specialist or other survey analyst to interpret evidence, evaluate alternative solutions, and make interactive updates. This approach applies less broadly; there are somewhat fewer than 0.2 million multiunit enterprises that own or operate 1.5 million establishments. However, maintenance of BR records for multiunits is disproportionately important because of their great economic impact, and this work consumes a majority of the BR's resources.

Update Mode and Data Source	Multiunit Enterprise	Multiunit Establishment	Single Unit	EIN Entity	Content of Updates	Frequency of Updates
INTERACTIVE UPDATES: BR Research and Correction Applications	X	X	X	X	Varies—most important for ownership/affiliation, multiunit composition & organization	Workdays[1]
IRS Business Master File			X	X	Identifying information, industrial classification, tax filing status	Complete—Annually (May) Updates—Monthly[1]
SSA Industry Codes (Business Births)			X	X	Industrial and other classifications	Monthly[1]
IRS Payroll Tax Returns			X	X	Mid-March employment, quarterly payroll measures for business employers, annual payroll measures for agricultural employers	Quarterly tax filings (annual for agricultural employers); data posted as received in weekly files from current IRS processing[1]
BLS Industry Codes			X	X	Industrial classification	Quarterly (temporarily suspended)
IRS Business Income Tax Returns			X	X	Receipts measures (vary by tax form), assets, industrial classification, name and address for nonemployer sole proprietors	Annual tax filings; data posted as received in weekly files (February - December) from current IRS processing[1]
BoC Company Organization Survey (Multiunits)	Indirect[2]	X[2]	X[3]	X[3]	Mid-March employment, annual and 1st quarter payroll, industrial classification, ownership/affiliation, multiunit composition & organization	Annual collection; data posted as received from survey processing (February - August)
BoC Economic Surveys	Indirect[2]	Indirect[2]	X	X	• Primarily industrial classification • Indirectly—ownership/affiliation, multiunit composition & organization	Annual, quarterly, and monthly collections; data posted as received
BoC Imputation Procedures for BR Data	Indirect[4]	X	X	X	Mid-March employment, quarterly payroll for EIN entities/ single units, annual and 1st quarter payroll for multiunit establishments	Annual: • July for EIN entities/ single units • September for multiunit establishments
BoC Economic Census	Indirect[2]	X[2]	X	X	Mid-March employment, annual and 1st quarter payroll for multiunit establishments; industrial classification; physical location; the BR's economic census content; ownership/affiliation, multiunit composition & organization	Every 5 years—ownership/affiliation, multiunit composition & organization, multiunit establishment data updated throughout census processing; final results posted about 18 months after end of reference year

[1] There is a brief lapse in updating during September for annual BR closeout operations.
[2] This source identifies changes to ownership/affiliation and multiunit composition & organization, but BR updates based on this information generally are made through interactive applications.
[3] The COS identifies affiliations between EIN entities and multi-establishment enterprises.
[4] Enterprises are affected indirectly, as payroll and employment for establishments are summed to produce enterprise totals.

Table 2. Variations in BR Updates: Mode, Source, Type of Unit, Content, and Frequency

Quality Assurance

15.   The Census Bureau has implemented procedures described below to assure the quality of BR data inputs and processes. Expansion and improvement of these procedures is an important goal, and the Census Bureau will proceed with enhancements as resources permit.

15.1 **Administrative Records**—Quality assurance for the BR's administrative records inputs is a two-stage process. The first stage evaluates aggregate data in order to identify global errors that may warrant rejection of a whole file. Specifically, it tabulates distributions of the file's variables and compares those distributions to standards based on levels and trends from three previous years' data. If the comparisons identify items that fail to meet these standards, the administrative records staff investigates those discrepancies and resolves them, usually by obtaining a corrected file or a reasonable explanation of the data from the source agency. The second stage uses record-by-record edits to identify reporting or processing errors that may affect individual tax reporting units (EINs) and individual variables recorded for them. Done as part of the process that updates the BR, the edits compare each input variable to corresponding BR data for prior periods and identify extreme variations as probable errors. When such a probable error is found, the editing process will not allow an BR update for the suspect variable; instead, it marks the input record for review. The administrative records staff monitors the edit results and notifies the source agency if systematic errors are found. Later, BR imputation procedures supply estimates for items that are blank because administrative data edits would not allow an update based on a suspect value (probable processing or reporting error).

15.2 **Census Bureau Collections**—The Census Bureau routinely applies quality assurance procedures to data collection and processing for the COS, the economic censuses, and other economic surveys. For straightforward operations like data entry, these procedures include independent verification of samples drawn from work lots and subsequent decisions to accept or reject (rework) those lots based on error rates estimated from the samples. For less deterministic operations, quality assurance procedures monitor work samples for acceptability and provide feedback needed to improve staff skills and resolve process errors. Additionally, data from Census Bureau collections are subjected to a variety of microdata edits and macrodata analyses designed to detect and resolve reporting errors as the information is prepared for use in statistical products. Editing and analysis for the annual County Business Patterns series are particularly important in this regard, as this statistical product presents tabulations of BR data, and the program's data corrections are incorporated directly.

Because they play a uniquely important role in maintaining BR coverage for multi-establishment enterprises, the COS and the multiunit segment of the economic censuses incorporate specialized "completeness and coverage" edits. These edits evaluate data for each enterprise as a whole, for each affiliated EIN entity, and for each owned/operated establishment in order to identify extreme variation as compared to administrative (tax) data for the current period and to corresponding COS or census data for prior periods. The aim is to identify coverage errors, such as omissions and duplications, as well as more routine errors in reporting of establishments' employment and payroll. Apparent problems are referred to a staff of COS specialists who use BR interactive applications and other resources to investigate discrepancies and correct errors.

Batch processes that update the BR based on census and survey feedback also monitor those updates and enforce data integrity rules. Each variable is examined for validity and consistency by procedures that precede the updating process, by edits carried out as part of the updating process, or by a combination of these approaches.

15.3 **Interactive Updates**—Interactive BR updates are carefully controlled by applications that assign Census Bureau employees to user classes with varying levels of access and update authority (for example, some user classes cannot make updates that affect multiunit composition or organization). Automated validity and consistency edits enforce data integrity rules for updates by all users; further, work samples may be subjected to strict independent verification or less structured quality monitoring, depending on user class and transaction type.

15.4   **Quality Profiles**—Lastly, the Census Bureau produces monthly tabulations that profile key BR attributes by broad industrial classification (2-digit NAICS) in order to detect questionable period-to-period variations in the level or distribution of primary measures. If these quality profiles reveal unusual changes, BR specialists review the profiles, trace changes to their source(s), evaluate the validity of underlying data, and take corrective action, if needed.

These quality profiles are an ongoing development project. The basic implementation done so far is very limited in the number of attributes it addresses and the amount of detail it provides; further, its emphasis is on resolving potential problems rather than estimating error rates. Still, it reveals useful information about fundamental BR characteristics that was not available a few years ago. The Census Bureau plans to build upon knowledge gained from these profiles and use them as a foundation for additional analyses that will examine a larger number of attributes for more detailed industrial and other classifications. The Census Bureau also hopes to refine detection techniques applied to these profiles so it can quickly identify and effectively resolve substantive problems in BR source data and maintenance processes.